Seminar Nasional Informatika 2011 (semnasIF 2011) *ISSN: 1979-2328*
UPN "Veteran" Yogyakarta, 2 Juli 2011

PERANCANGAN PROGRAM APLIKASI PERAMALAN BANJIR KANAL BARAT JAKARTA MENGGUNAKAN *AUTOREGRESI MULTIVARIANT*

Ngarap Im Manik
Jurusan Matematika FST BINUS University,
Jln.Kebon Jeruk Raya no.27 Jakarta Barat 11480, Indonesia
email : manik@binus.edu

Abstrak
Makalah ini membahas tentang perancangan program komputer yang dapat mengetahui gambaran karakteristik data ketinggian permukaan air di pintu air sungai Manggarai, mengetahui variabel mana yang paling berpengaruh terhadap ketinggian permukaan air dan upaya pencarian model peramalan banjir yang lebih baik dengan menggunakan metode autoregresi multivariant. Hasil penelitian ini dapat membantu petugas pintu air dalam memberi peringatan dini dan antisipasi upaya pencegahan dan penanggulangan banjir. Model persamaan peramalan yang didapat adalah $Y_t = 109,.7828 + 0,9291\ CH_{t\text{-}6} - 24,484\ T_{t\text{-}2} - 0,06245\ PM_{t\text{-}2} + 1,4706\ KB_{t\text{-}2}$ *di mana temperatur dan ketinggian permukaan air merupakan variabel yang memiliki hubungan yang paling kuat. Variabel ini memiliki hubungan secara negatif yang berarti ketika temperatur turun maka nilai ketinggian permukaan air akan naik. Koefisien determinasi memiliki nilai sebesar* $R^2 = 0.4056$ *dan statistik Durbin Watson sebesar* $DW = 0.7429$*.*

***Kata Kunci** : Peramalan banjir, autoregresi multivariant, kanal barat Jakarta,*

1. PENDAHULUAN

Banjir biasanya disebabkan oleh beberapa faktor. Salah satu faktor yang penting adalah curah hujan. Banjir di Jakarta yang menyebabkan lumpuhnya ibu kota Negara tahun 2007 silam disebabkan karena meningkatnya intensitas curah hujan pada bulan Oktober hingga bulan Maret yang mencapai puncaknya pada bulan Febuari. Pada bulan-bulan ini berjuta-juta gallon air di tumpahkan ke Jakarta melalui sistem drainase yang ada. Namun karena buruknya sistem drainase di Jakarta maka air yang mengalir tidak dapat tersalurkan dengan baik sehingga terjadilah luapan air. Namun dalam pembahasan ini tidak membahas tentang kerusakan sistem drainase. Tetapi kita akan melihat bahwa perubahan curah hujan pada bulan-bulan tertentu akan menyebabkan perubahan tinggi permukaan air serta debitnya. Perubahan yang teratur dalam jangka waktu tertentu berarti menghasilkan pola data *time series.* Jika data *times series* bisa kita dapatkan maka dapat dilakukan peramalan pada data-data yang telah lalu untuk memperkirakan bagaimana pola datanya di masa yang akan datang. Selanjutnya karena permasalahan pada penelitian ini begitu luas maka perlu diberikan batasan-batasan permasalahan yang dibahas, yaitu : Daerah yang diuji terbatas untuk daerah pintu air Manggarai Kanal Barat Jakarta yang meliputi daerah Kapuk Muara dan Pluit dengan melewati sepanjang jalan Sultan Agung (Stasiun Dukuh), jalan Galunggung, KH Margono Djojohadikoesoemo, Petamburan, Stasiun Tanah Abang. Lalu membelah jalan Kyai Caringin-Tomang Raya dan KH Hasyim Asy'ari-Kyai Tapa menuju Angke. Kemudian data aliran air didapatkan dari Pemberdayaan Sumber Daya Air dan Pantai dalam bentuk data debit air dan tinggi permukaan air dalam periode waktu tiap jam untuk ketinggian permukaan dan data harian dari stasiun Meteorologi dan geofisika Bogor.

Sampai saat ini PSDA (Pemberdayaan Sumber Daya Air dan Pantai) dikelola oleh badan pemerintah yang berkewenangan dalam mengendalikan banjir melalui sistem pintu air yang masih belum mempunyai sistem peramalan. Aplikasi sistem peramalan ini nantinya tidak hanya digunakan untuk Jakarta namun juga bisa di gunakan di daerah-daerah lain. Dengan menggunakan metode analisis deret berkala multivariant maka penulis mencoba untuk membuat suatu pemodelan yang aktual tentang hubungan antara curah hujan, temperatur, lama penyinaran matahari, kelembaban nisbi dan ketinggian permukaan air sungai. Sungai yang diteliti kali ini adalah sungai Manggarai. alasan penulis mengambil sungai Manggarai karena sungai ini mengendalikan aliran air di tengah-tengah kota Jakarta dan juga sebagai palang pintu terakhir sebelum air masuk ke Istana Negara. Untuk pembentukan model diambil dari seluruh data yang sudah didapat, mulai dari tahap explorasi data sampai dengan peramalan yang di harapkan.

2. TINJAUAN PUSTAKA

2.1 Banjir Kanal Barat

Banjir yang kita rasakan terutama bila memasuki awal tahun biasanya terjadi pada bulan-bulan Januari dan Febuari, pada dasarnya banjir karena kiriman air dari Bogor. Hujan di Bogor yang cukup deras akan menyebabkan air di sungai-sungai Jakarta meluap. Salah satunya adalah sungai yang diamati oleh peneliti.

Sungai Manggarai mengendalikan banjir di Kanal Barat yang akan berpengaruh pada daerah lain seperti Kapuk Muara dan Pluit dengan melewati sepanjang jalan Sultan Agung (Stasiun Dukuh), jalan Galunggung, KH Margono Djojohadikoesoemo, Petamburan, Stasiun Tanah Abang. Lalu membelah jalan Kyai Caringin-Tomang Raya dan KH Hasyim Asy'ari-Kyai Tapa menuju Angke. Perlu di ketahui bahwa sungai Manggarai juga mengendalikan aliran sungai ke Istana Negara.

Curah hujan yang terjadi di Bogor tidak sepenuhnya langsung mengalir ke Jakarta. Diperlukan waktu untuk mengalami peresapan terlebih dahulu. Sedangkan tanah tidak akan menyerap terlalu cepat jika tanah masih menampung sisa-sisa air dari waktu yang lalu. Dalam hal ini tanah sering mengalami kejenuhan sehingga tidak dapat mengalirkan air secara langsung yang diterima dari air hujan. Dengan adanya teori ini maka penulis menyimpulkan bahwa data curah hujan tidak bisa langsung dianalisis bersama-sama dengan ketinggian permukaan air sungai. Diperlukan penjumlahan tiap variabel hujan tersebut. Sebagai contoh disini penulis menjumlahkan variabel hujan selama seminggu, untuk mendapatkan model yang lebih mewakili keadaan sebenarnya.

Adapun variabel yang di pakai dalam penelitian ini adalah variabel ketinggian permuakaan air sungai Manggarai sebagai variabel yang *dependent* dan variabel curah hujan di Bogor sebagai variabel *independent.* Karena kita semua bisa mengasumsikan bahwa ketinggian permukaan air di sungai Manggarai sangat bergantung pada curah hujan di Bogor. Oleh karena itu semuanya akan saling berhubungan.

2.2 Ketepatan Metode Peramalan

Dalam banyak situasi peramalan, ketepatan dipandang sebagai kriteria penolakan untuk memilih suatu metode peramalan, dalam banyak hal, kata ketepatan (*acuracy*)" menunjuk ke "kebaikan suai", yang pada akhirnya penunjukan seberapa jauh model peramalan tersebut mampu mereproduksi data yang telah diketahui. Dalam pemodelan eksplanatoris (kausal), ukuran kebaikan suai cukup menonjol. Dalam pemodelan deret-berkala, sebagian data yang diketahui dapat digunakan untuk meramalkan sisa data berikut sehingga memungkinkan orang untuk mempelajari ketepatan ramalan secara lebih langsung. Bagi pemakai ramalan, ketetapan ramalan yang akan datang adalah yang paling penting. Bagi pembuat model, kebaikan suai model untuk fakta (kuantitatif dan kualitatif) yang diketahui harus diperhatikan.

2.2.1 Koefisien Korelasi

Seringkali terjadi bahwa dua variabel dikaitkan satu sama lain, walaupun mungkin tidak selalu benar bahwa nilai suatu variabel bergantung pada, atau disebabkan oleh perubahan nilai variabel yang lain. Pada setiap kejadian, suatu hubungan data dinyatakan dengan perhitungan korelasi antara dua variabel. Koefisien korelasi r adalah suatu ukuran asosiasi (linear) relatif antara dua variabel. Ia dapat bervariasi dari 0 (yang menunjukan tidak ada korelasi) hingga ±1 (yang menunjukan korelasi sempurna). Jika korelasi lebih besar dari 0, dua variabel dikatakan berkorelasi positif dan jika kurang dari 0 dikatakan korelasi negatif. Koefisien korelasi memegang peranan penting dalam analisis data multivariant (yaitu apabila yang terlibat dua variabel atau lebih) dan mempunyai kaitan erat dengan analisis regresi.

2.2.2 Perhitungan Koefisien Korelasi

Untuk menghitung korelasi antara dua variabel X dan Y yang dinotasikan sebagai r_{XY} untuk n pasangan observasi (Xi, Yi), i = 1,2,…,n, rumus-rumus berikut adalah relevan: Nilai tengah X

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{1}$$

Nilai tengah Y

$$\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i \tag{2}$$

Kovarians antara X dan Y

$$Cov_{XY} = \frac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}) \tag{3}$$

Varians X

$$Cov_{XX} = Var_X = \frac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X})^2 = S_X^2 \tag{4}$$

Varians Y

$$Cov_{YY} = Var_Y = \frac{1}{n}\sum_{i=1}^{n} (Y_i - \overline{Y})^2 = S_Y^2 \tag{5}$$

Korelasi antara *X* dan *Y*

$$r_{XY} = \frac{Cov_{XY}}{\sqrt{Cov_{XX} Cov_{YY}}} = \frac{Cov_{XY}}{S_X S_Y} \qquad (6)$$

Dimana $S_x = \sqrt{Cov_{XX}}$ dan $S_Y = \sqrt{Cov_{YY}}$ adalah deviasi standar *X* dan *Y*.

Perhatikan bahwa pada rumus-rumus tersebut, semua penjumlahan dibagi dengan *n* dan bukan *n* – 1 pada rumus kovarians dan ragam maka tidak akan mengubah rumus korelasi pada persamaan di atas.

Rumus lain untuk menghitung koefisien korelasi adalah sebagai berikut:

$$r_{XY} = \frac{n\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{n\Sigma X^2 - (\Sigma X)^2}\sqrt{n\Sigma Y^2 - (\Sigma Y)^2}} \qquad (7)$$

Keuntungan lain dari rumus ini adalah untuk setiap himpunan observasi berpasangan, hanya lima penjumlahan dasar yang harus di hitung yaitu

$$\Sigma X, \Sigma Y, \Sigma X^2, \Sigma Y^2, dan\ \Sigma XY$$

2.3 Regresi

Analisis regresi adalah metologi statistik yang berfungsi untuk memperkirakan hubungan dari dua atau lebih variabel kuantitatif dengan demikian salah satu variabelnya dapat diramalkan dari variabel lainya. (**Netter, Kutner Nachtseim, Wasserman, 1996, p3**). Analisis regresi dikembangakan pertama kali oleh Sir Grancis Galton di akhir abad 19. Galton telah mempelajari hubungan antara ketinggian orang tua dengan anak-anak mereka dan menemukan bawa ketinggian anak baik yang pendek maupun yang tinggi terlihat mengarah ke arah rata-rata kelompok. Dia menyimpulkan bahwa data ini bergerak mengikuti aturan ke arah kemunduran ke nilai tengah. Galton mengembangkan diskripsi matematika untuk kemunduran yang mengikuti aturan ini, yang sekarang kita kenal dengan model regresi. Terdapat dua jenis hubungan antar variabel dalam regresi yaitu hubungan fungsional dan hubungan statistikal.

Hubungan fungsional antar dua variabel dapat di ekspresikan dalam formula matematika. Jika X didefinisikan sebagai variabel bebas dan Y sebagai variabel tak bebas, maka hubungan fungsionalnya adalah seperti berikut : $Y = f(X)$ Hubungan secara statistikal, hampir sama seperti hubungan fungsional tetapi bukan hubungan secara sempurna. Di dalam observasi secara umum hubungan statistikal tidak berpengaruh secara langsung terhadap kurva hubungan. Analisis regresi memiliki tiga tujuaan: (1) deskripsi, (2) kontrol, (3) peramalan. (**Netter, Kutner Nachtseim, Wasserman, 1996, p9**). Deskripsi karena analisis regresi mampu memberikan gambaran hubungan antar variabel. Kontrol berguna jika kita ingin mengetahui sebarapa besar nilai variabel bebas jika kita ingin nilai variabel tak bebasnya sebesar nilai tertentu. Selain hal tersebut model regresi juga dapat digunakan sebagai penduga variabel tak bebas dengan variabel-variabel bebasnya. Secara umum bentuk persamaan regresi adalah sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad (8)$$

Dimana :

Y_i adalah nilai respon dari variabel ke i ; β_0 dan β_1 adalah parameter

X_i adalah nilai variabel prediktor ke i ; ε_i adalah *random error* dgn rata-rata E{ ε_i }=0 dan ragam σ^2

2.3.1 Regresi Berganda

Regresi berganda adalah metode yang sesuai untuk analisis ketika problem penelitian melibatkan satu variabel tak bebas yang di asumsikan mempunyai hubungan dengan dua atau lebih variabel bebas (**Hair, Anderson, Tatham, Black, 1998, p14**). Tujuan dari regresi berganda adalah untuk memprediksi perubahan dari variabel tak bebas yang berespon terhadap perubahan variabel bebas. Tujuan ini dapat dicapai melalui aturan statistik kuadrat terkecil.

Model regresi berganda (**Makridakis, 1999, p278**)

$$Y = \beta_0 + \beta_1 X_1 + ... + \beta_k X_k + \varepsilon_i \qquad (9)$$

Dimana :

$\beta_0, \beta_1, \beta_2, ..., \beta_k$ adalah parameter tetap, dan $X_1, X_2, ..., X_k$ konstata variabel bebas,

ε adalah suatu variabel random yang menyebar secara normal di sekitar nol (nilai tengah ε) dan mempunyai suatu ragam V_ε.

Perhatikan bahwa bentuk model regesi pada persamaan (9) adalah linear pada koefisienya. Pangkat dari setiap koefisien β adalah 1 (linear) dan ini berati bahwa taksiran koefisien β dapat diperoleh secara efisien dengan menggunakan metode kuadrat terkecil (*least square = LS method*). Jika terdapat dua variabel, berarti Y dipetakan dalam sebuah bidang. (bidang yang dibentuk dari dua buah variabel X). Jika terdapat lebih dari dua variabel bebas maka kita katakan bawa Y dipetakan ke dalam sebuah bidang *hyperplane* (yang berarti suatu permukaan berdimensi lebih tinggi).

Dalam praktiknya, tugas pemodelan regresi adalah untuk menaksir parameter yang tidak diketahui pada model, yaitu $\beta_0, \beta_1, \beta_2, ..., \beta_k$ dan V_ε. Dari himpunan data yang diketahui. Prosedur LS dapat diterapkan untuk menentukan $\beta_0, \beta_1, \beta_2, ..., \beta_k$ dan menaksir V_ε. Akar kuadrat dari taksiran terakhir ini sering kali disebut: galat standar taksiran (*standard error of estimate*). Dengan demikian bentuk pragmatis model regesi secara statistik adalah sebagai berikut :

$$\hat{Y} = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + ... + \beta_k X_{ki} + e_i \tag{10}$$

Dimana :

$\beta_0, \beta_1, \beta_2, ..., \beta_k$ adalah penaksir LS dari $\beta_0, \beta_1, \beta_2, ..., \beta_k$, dan semuanya adalah variabel acak, dengan sebaran bersama normal.

$X_1, X_2, ..., X_k$ konstata variabel bebas,

$e_i (i = 1, 2, ..., N)$ adalah suatu nilai galat taksiran, untuk pengamatan ke *i*, dan diasumsikan merupakan sampel independent dari suatu sebaran binomial.

2.3.2 Autoregresi

Secara umum proses autoregresi (AR) orde ke-p, mempunyai bentuk model sebagai berikut:

ARIMA (*p*,0,0)

$$X_t = \mu' + \beta_1 X_{t-1} + \beta_2 X_{t-2} + ... + \beta_p X_{t-p} + e_t \tag{11}$$

Di mana : μ = nilai konstan ; β_j = parameter autoregresi ke-j ; e_t = nilai galat pada saat t

Model autogresi sering disebut juga model ARIMA (*p*,0,0). Karena angka pertama pada model ARIMA melambangkan autoregresi atau sering disebut AR, sedangkan untuk angka ketiga kita kenal sebagai ordo rata-rata bergerak atau sering kita sebut sebagai MA (*moving average*). Sedangkan jika ketiga angkanya bukan nol maka hal tersebut menunjukan model campuran antara autoregresi dan rata-rata bergerak.

2.4 Rekayasa Perangkat Lunak

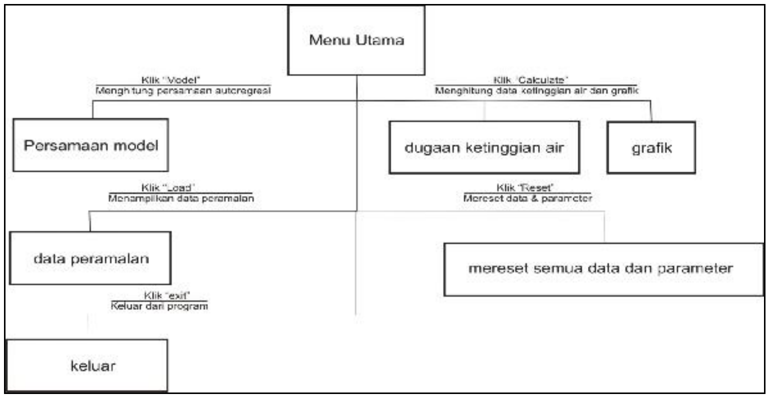
Model rekayasa piranti lunak yang dipakai penulis adalah model sekuensi linear. Model ini disebut juga model"air terjun" (*waterfall*).Model ini merupakan sebuah pendekatan kepada perkembangan perangkat lunak yang sistematik dan sekuensial yang muali pada tingkat dan kemajuan sistem pada seluruh analisis, desain, koding, pengujian dan pemeliharaan :

3. METODE PENELITIAN

Sesuai dengan judul makalah ini maka metode penelitian lebih difokuskan dalam perancangan programnya. Perancangan program aplikasi *autoregresif multivariate* bertujuan untuk membentuk model *autoregresi* (AR) dengan empat buah variabel yang sudah ditetapkan. Dari program ini dapat dilihat pola-pola yang terbentuk. Dari mulai pola data sebenarnya, data eksporasi, data peramalan dan perbandingan antara data sebenarnya dengan data peramalan. Dengan melihat tampilan data berupa grafik diharapkan agar lebih mudah dipahami sekalipun untuk orang awam. Adapun metodologi yang dibuat terdiri dari tiga bagian yaitu : Perancangan diagram STD, Perancangan *flowchart*, dan perancangan layar.

3.1 Perancangan Diagram Transisi (STD)

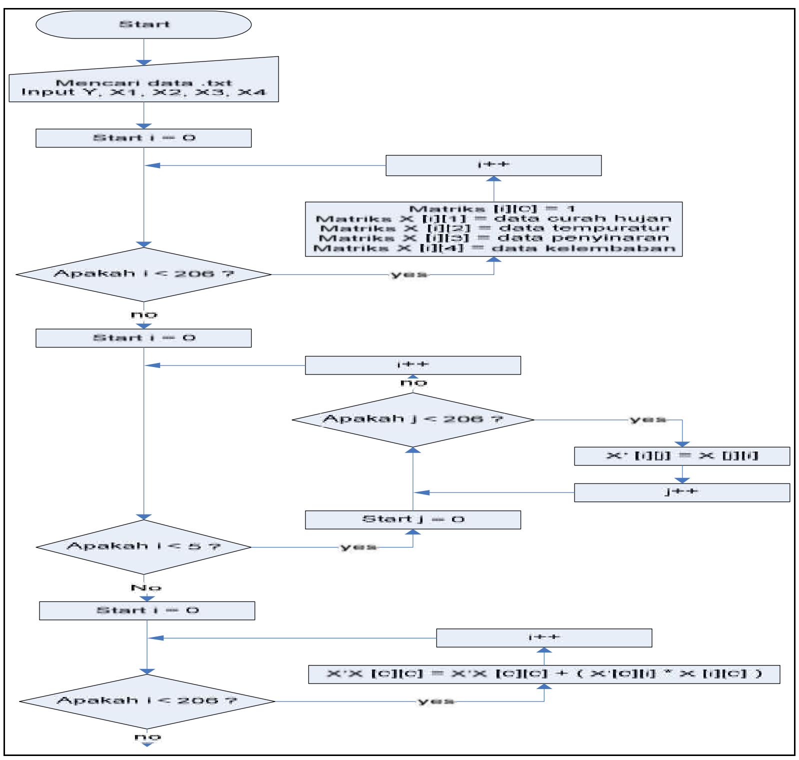
Diagram transisi digunakan untuk memberi gambaran secara menyeluruh tentang cara kerja suatu sistem aplikasi. Dengan menggambarkan keadaan (*state*) dan aksi (*event*) yang menyebabkan sistem tersebut berpindah keadaan, seperti yang ditunjukkan dalam gambar 1.



Gambar 1. Diagram Transisi Program

3.2 Perancangan Diagram Alir (*Flowchart*)

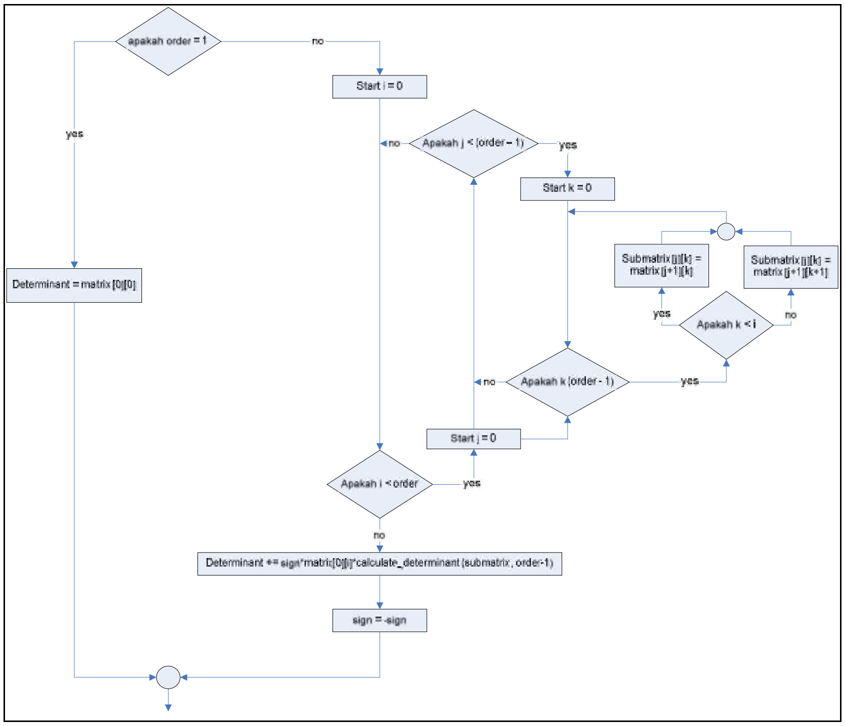
Diagram alir ini dibuat untuk menggambarkan interaksi antara perhitungan metode *autoregresif multivariate* dengan sistem aplikasi yang dibuat. Dengan adanya flowchart ini maka kita dapat lebih jelas melihat bagaimana perhitungan yang dilakukan oleh program (proses perhitungan yang biasanya tidak ditampilkan pada program aplikasi). Hal ini dalam dilihat pada gambar 2.



(**)

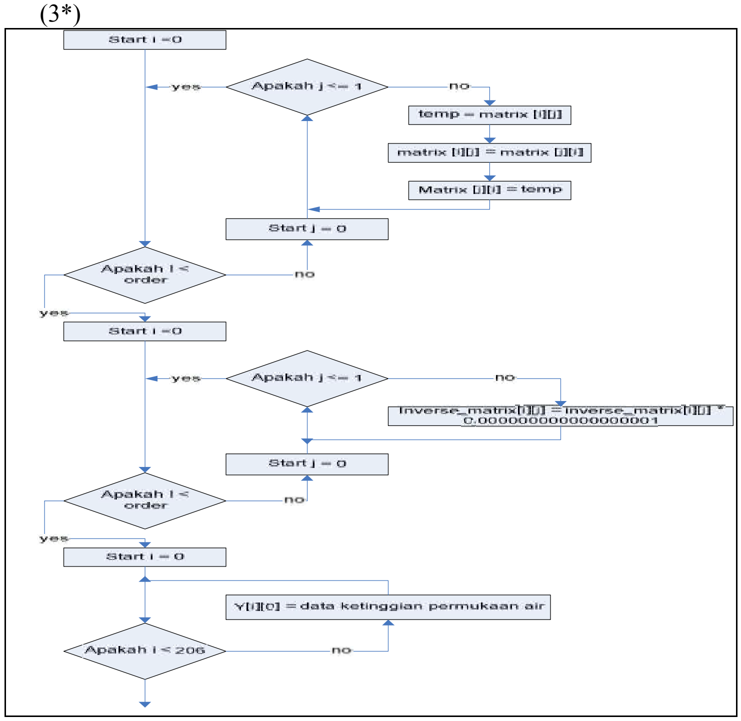
Mengulang kembali ke Start i = 0 pada X'X [0][1], X'X[1][1] sampai dengan X'X[4][4]. Kemudian dilanjutkan ke diagram berikut ini

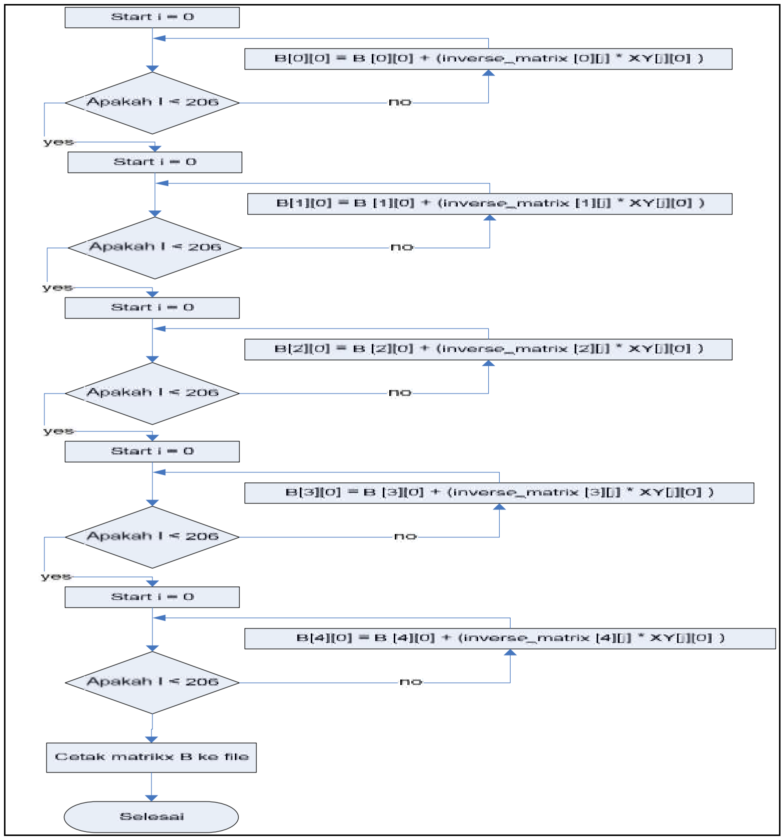
(**)


apakah order = 1
no
yes
Start i = 0
Apakah j < (order - 1)
no
yes
Start k = 0
Submatrix [j][k] = matrix [j+1][k]
Submatrix [j][k] = matrix [j+1][k+1]
Apakah k < i
yes
no
Determinant = matrix [0][0]
Apakah k (order - 1)
no
yes
Start j = 0
Apakah i < order
yes
no
Determinant += sign*matriz[0][i]*calculate_determinant (submatrix, order-1)
sign = -sign

(3*)

(3*)

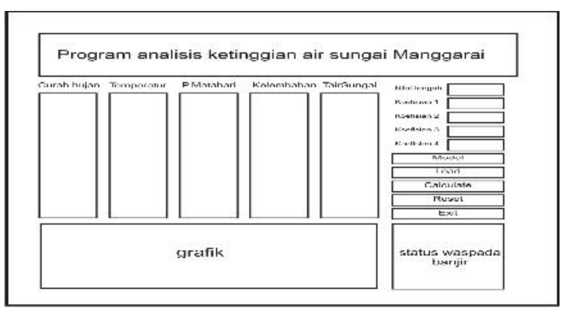

Start i = 0
Apakah j <= 1
yes
no
temp = matrix [i][j]
matrix [i][j] = matrix [j][i]
Matrix [j][i] = temp
Start j = 0
Apakah i < order
no
yes
Start i = 0
Apakah j <= 1
yes
no
Inverse_matrix[i][j] = inverse_matrix[i][j] * 0.0000000000000000001
Start j = 0
Apakah i < order
no
yes
Start i = 0
Y[i][0] = data ketinggian permukaan air
Apakah i < 206
no



Gambar 2. Flowchart Autoregresi

3.3 Perancangan Layar

Perancangan layar adalah gambaran layar muka pengguna (*user interface*) yang diberikan oleh program. Adapun rancangan program deteksi dini banjir kanal barat dengan menggunakan program metode *autoregresi multivariate* adalah sebagai berikut :

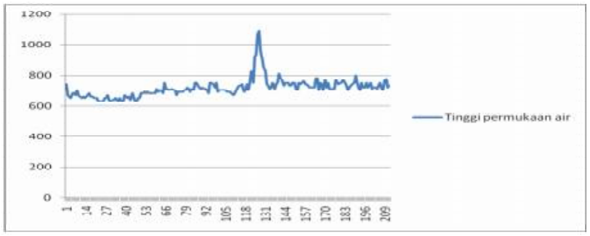


Gambar 3. Rancangan Layar

4. HASIL & PEMBAHASAN

Hasil pengumpulan data

Berdasarkan catatan harian pintu air Manggarai maka diperoleh jumlah populasi data ketinggian permukaan sungai selama bulan Oktober 2009 sampai April 2010 adalah sebanyak 5112 data. Data ini tidak dipakai semuanya. Dalam kenyataannya penulis akan meringkas ke dalam bentuk harian. Yang dapat kita hitung bahwa dari bulan Oktober sampai April ada 212 hari demikian pula banyaknya data yang akan diteliti. Sebenarnya data yang di dapat dari catatan harian pintu air Manggarai tidak hanya data ketinggian permukaan air saja. Tetapi ada juga debit air, sisa, dan tingginya pembukaan pintu air namun untuk variabel yang digunakan hanyalah ketinggian permukaan air hal ini berdasarkan wawancara langsung ke petugas penjaga pintu air, dan juga karena pencatatan debit air, sisa maupun pembukaan pintu adalah berdasarkan ketinggian permukaan air bukan berdasarkan alat ukur khusus. Data dimaksud ditunjukkan dalam gambar 4.



Gambar 4. Grafik Ketinggian Permukaan Air Manggarai Okt 2009- Apr 2010

Dapat kita lihat bahwa data di atas memiliki hubungan linear antara waktu dan ketinggiannya. Dan tampaknya juga memiliki pola data musiman dengan data pencilan atas di bagian tengahnya. Pencilan tersebut merupakan data banjir bandang yang ternyata sungai Manggarai mencapai titik tertinggi pada 4 Febuari 2009 yaitu 1090 meter dari dasar sungai. Tetapi kita tidak akan membahas lebih dalam mengenai hal itu. Data tersebut mempunyai rata-rata sebesar 713.94 dan standart deviasi sebesar 56.51.

4.1.1 Menghilangkan data pencilan

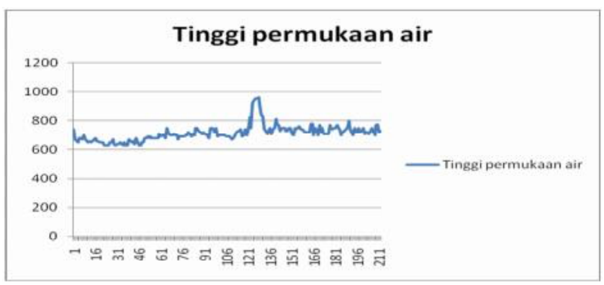
Dalam analisis kolerasi dan regesi secara umum data pencilan atau sering disebut dengan *outlayer* bisa mengakibatkan nilai-nilai koefisien kolerasi dan regresi tertarik/*skewed* kearah data pencilan tersebut. Dengan adanya *skewness* ini akan menyebabkan peramalan menjadi tidak valid lagi karena akan menjauhi nilai rata-ratanya yang sebenarnya. Oleh sebab itu dalam penelitian ini data pencilan akan dihilangkan.

Data sekitar hari ke-131 pada grafik, terlihat bahwa terdapat data pencilan atas. Untuk menghilangkannya maka digunakanlah rata-rata dari sebelum dan sesudah pencilan itu terjadi dengan demikian dapat memperbaiki keakuratan model yang di dapat.

Dapat kita lihat rumus tepatnya adalah sebagai berikut:

$$\hat{Y} = \frac{Y_{126} + Y_{127}}{2}$$

Dengan pendugaan diatas maka kita bisa melihat kembali grafiknya menjadi seperti yang ditunjukkan pada gambar 5.



Gambar 5. Grafik Ketinggian Permukaan Air Setelah Menghilangkan Pencilan

Dari grafik masih dapat kita lihat bawah sekitar hari ke 121-136 terdapat data dengan pola fluktuasi yang cukup ekstrim. Namun data ini tidak perlu dihilangkan karena merupakan bagian dari pola data itu sendiri bukan merupakan data pencilan.

4.1.2 Korelasi antar variabel

Sebelum membentuk sebuah model peramalan, hal yang kita perlukan adalah mengetahui besarnya hubungan antar variabel baik itu secara positif maupun secara negatif. Dengan mengetahui besarnya hubungan antara variabel ini maka kita bisa mengetahui variabel mana yang terbaik untuk dimasukan ke dalam model. Tentu saja dengan adanya variabel yang baik akan membuat model peramalan kita mempunyai ragam yang cukup bervariasi dan juga *error* yang lebih kecil.

Korelasi biasanya berguna untuk memantau besarnya hubungan antar variabel. Namun di penelitian ini kita akan melihat hubungan antar variabel dengan penambahan waktu jeda (*time lag*). Waktu jeda disini dimasukan karena pemodelan autoregesi yang akan kita pakai akan menggunakan waktu jeda. Untuk lebih jelas kita bisa melihat modelnya secara langsung :

$$Y_t = \mu' + \beta_1 X_{t-1} + \beta_2 X_{t-2} + ... + \beta_p X_{t-p} + e_t$$

Dapat kita lihat bahwa nilai $X_{t-1}, X_{t-2}, ..., X_{t-p}$ menunjukan adanya waktu jeda (*time lag*). Selain itu tanpa waktu jeda kita tidak akan bisa meramal kejadian yang akan datang dan jika itu terjadi maka kita hanya bisa menduga nilai variabel *Y* (variabel tak bebas) dari variabel *X* (variabel bebas) lainnya, tidak akan memungkinkan kita menduga nilai *Y* dari variabel lain sebelum saat itu tanpa waktu jeda.

Jeda yang akan kita ambil dibatasi dengan jangka waktu tujuh hari atau seminggu. Sehingga peramalan yang akan dilaksanakan tidak akan lebih dari waktu jeda ke tujuh. Maka secara garis besar pengujian korelasinya akan seperti berikut :

Ketinggian air saat ini (Yt) dipengaruhi oleh : X_1 = Curah hujan ; X_2 = Temperatur; X_3 = Lama penyinaran matahari; X_4 = Kelembaban nisbi

Keempat variabel tersebut mempunyai time lag satu hingga tujuh sebelum dilakukan pengujian korelasi.

Korelasi dengan *time lag* satu

Tabel 1. Koefisien Korelasi dengan *Time Lag* Satu

	Tinggi Permukaan	Curah hujan	Temperatur	Lama Penyinaran	Kelembaban Nisbi
Tinggi Permukaan	1				
Curah hujan	0.123432128	1			
Temperatur	-0.546510491	-0.182879557	1		
Lama Penyinaran	-0.368289391	-0.099056552	0.637881557	1	
Kelembaban Nisbi	0.38921876	0.175795539	-0.49837388	-0.503480284	1

Dari tabel 1 dapat dilihat bahwa secara keseluruhan hubungan antara data ketinggian permukaan air dan data klimatologi (curah hujan, tempratur, lama penyinaran matahari, dan kelembaban nisbi) adalah seperti yang ditampilkan pada tabel 2.

Tabel 2. Hubungan Korelasi Antar Variabel

Y	X	Hubungan
Ketinggian permukaan air	Curah hujan (°C)	Positif
Ketinggian permukaan air	Temparatur (mm)	Negatif
Ketinggian permukaan air	Lama penyinaran (%)	Negatif
Ketinggian permukaan air	Kelembaban nisbi (%)	Positif

Hubungan yang positif menunjukan bahwa semakin besar nilai variabel yang satu akan semakin besar pula nilai variabel yang dipengaruhinya. Sebagai contoh semakin tinggi curah hujan yang terjadi di Bogor maka semakin tinggi pula permukaan air yang ada di pintu air Manggarai. Dan sebaliknya ketika temperatur menurun maka ketinggian pemukaan air akan naik. Karena jika di Bogor sering terjadi hujan maka secara otomatis suhu udara akan turun. Demikian juga dengan lama penyinaran matahari, jika matahari semakin jarang bersinar maka kemungkinan permukaan air akan naik menjadi besar. Untuk melihat sebarapa besar kuatnya pengaruh yang diberikan variabel-variabel tersebut kita bisa melihatnya di tabel 1 sebelumnya. Tampaklah di situ bahwa variabel yang paling besar memberi pengaruh terhadap permukaan air sungai adalah variabel temperatur yang mempunyai nilai sebesar -0.5465 yang berarti bahwa temperatur memiliki pengaruh positif setengah terhadap tinggi permukaan.

Korelasi dengan *time lag* dua

Tabel 3. Koefisien Korelasi dengan *Time Lag* Dua

	Tinggi Permukaan	Curah hujan	Temperatur	Lama Penyinaran	Kelembaban Nisbi
Tinggi Permukaan	1				
Curah hujan	0.177099558	1			
Temperatur	-0.555111832	-0.181881503	1		
Lama Penyinaran	-0.414242339	-0.094003887	0.639847352	1	
Kelembaban Nisbi	0.391361555	0.169902671	-0.504753468	-0.491534222	1

Korelasi dengan *time lag* tiga

Tabel 4. Koefisien Korelasi dengan *Time Lag* Tiga

	Tinggi Permukaan	Curah hujan	Temperatur	Lama Penyinaran	Kelembaban Nisbi
Tinggi Permukaan	1				
Curah hujan	0.189271803	1			
Temperatur	-0.542313202	-0.181780115	1		
Lama Penyinaran	-0.358044486	-0.091119817	0.640907437	1	
Kelembaban Nisbi	0.352612347	0.165341555	-0.508348564	-0.487932559	1

Korelasi dengan *time lag* empat

Tabel 5. Koefisien Korelasi dengan *Time Lag* Empat

	Tinggi Permukaan	Curah hujan	Temperatur	Lama Penyinaran	Kelembaban Nisbi
Tinggi Permukaan	1				
Curah hujan	0.252129843	1			
Temperatur	-0.518445879	-0.180475121	1		
Lama Penyinaran	-0.331498777	-0.088950174	0.640372886	1	

Kelembaban Nisbi	0.325255579	0.161226934	-0.508033784	-0.485954175	1

Korelasi dengan *time lag* lima

Tabel 6. Koefisien Korelasi dengan *Time Lag* Lima

	Tinggi Permukaan	Curah hujan	Temperatur	Lama Penyinaran	Kelembaban Nisbi
Tinggi Permukaan	1				
Curah hujan	0.258669057	1			
Temperatur	-0.518401365	-0.175970892	1		
Lama Penyinaran	-0.318037509	-0.086623957	0.639766819	1	
Kelembaban Nisbi	0.314391907	0.154758927	-0.49552895	-0.486704728	1

Korelasi dengan *time lag* enam

Tabel 7. Koefisien Korelasi dengan *Time Lag* Enam

	Tinggi Permukaan	Curah hujan	Temperatur	Lama Penyinaran	Kelembaban Nisbi
Tinggi Permukaan	1				
Curah hujan	0.275144229	1			
Temperatur	-0.466315469	-0.17190314	1		
Lama Penyinaran	-0.255466606	-0.082986799	0.636579234	1	
Kelembaban Nisbi	0.262155518	0.148683088	-0.486242925	-0.481080267	1

Korelasi dengan time lag tujuh

Tabel 8. Koefisien Korelasi dengan *Time Lag* Tujuh

	Tinggi Permukaan	Curah hujan	Temperatur	Lama Penyinaran	Kelembaban Nisbi
Tinggi Permukaan	1				
Curah hujan	0.267882538	1			
Temperatur	-0.454865074	-0.16598566	1		
Lama Penyinaran	-0.274249768	-0.077457919	0.628468544	1	
Kelembaban Nisbi	0.258944215	0.140881288	-0.457485595	-0.467499136	1

Dari tabel-tabel di atas kita bisa mengetahui bahwa data curah hujan memiliki hubungan variabel terendah dari variabel lainya. Hal ini mungkin disebabkan karena data curah hujan terlalu banyak data yang bernilai nol. Nilai nol didapat karena tidak setiap hari kita akan mengalami hujan. Sehingga ketika tidak ada hujan tentu saja akan bernilai nol. Fenomena ini menyebabkan data curah hujan mempunyai ragam yang kecil sehingga menyebabkan hubungannya dengan variabel lain menjadi kurang bagus.

Data curah hujan sendiri terus naik sampai pada *time lag* yang ke enam dan turun pada *time lag* tujuh. Hal ini mendukung teori tentang titik jenuh tanah adalah benar. Bahwa ketika hujan turun maka tanah tidak dapat langsung menyerapnya. Perlu ada beberapa waktu sampai tanah mampu menyerap dan mengalirkannya ke sungai. Kemungkinan besar waktu tersebut adalah enam hari dengan melihat fenomena yang terjadi pada percobaan korelasi di atas. Untuk data-data lain semuanya tampak wajar dengan semakin menurunnya nilai variabel temperatur, lama penyinaran matahari, dan kelembaban nisbi dengan bertambahnya lag. Nilai terbaik mereka adalah saat mereka berada di lag kedua.

4.1.3 Autokorelasi

Setelah mengetahui besarnya hubungan korelasi antar variabel dengan setiap *time lag* berarti kita siap untuk membentuk model dengan persamaan autoregresi. Model yang akan kita pakai adalah model AR (4,0,0). AR (4,0,0) akan mewakili tiap variabel dengan *time lag* tertentu. Variabel yang terwakili urutannya adalah sebagai berikut yaitu curah hujan, temperatur, lamanya penyinaran matahari, dan kelembaban nisbi. Untuk mengetahui *time lag* terbaik pada tiap persamaan itulah maka kita menghitung variabel koelerasi seperti yang kita lakukan di sub bab sebelumnya. Tabel 9 menunjukan waktu terbaik tiap-tiap variabel.

Tabel 9. Waktu Terbaik Tiap Variabel

Variabel	Waktu terbaik
Curah hujan (^oC)	Time lag enam
Temparatur (mm)	Time lag dua
Lama penyinaran (%)	Time lag dua
Kelembaban nisbi (%)	Time lag dua

Mengikuti waktu terbaik yang telah disajikan tabel di atas maka pemodelan regresinya menjadi :

AR (4,0,0) $Y_t = \mu' + \beta_1 X_{t-6} + \beta_2 X_{t-2} + \beta_3 X_{t-2} + \beta_4 X_{t-2}$

Setelah mendapatkan model di atas maka kita bisa mulai menghitung setiap nilai koefisiennya dengan menggunakan cara regresi berganda. Dan hasil yang di dapat dilihat pada tabel 10 berikut :

Tabel 10. Hasil Perhitungan Regresi

	Coefficients
Intercept	1092.782851
X Variabel 1	0.929105052
X Variabel 2	-24.48475667
X Variabel 3	-0.062451946
X Variabel 4	1.470628098

Persamaannya menjadi :

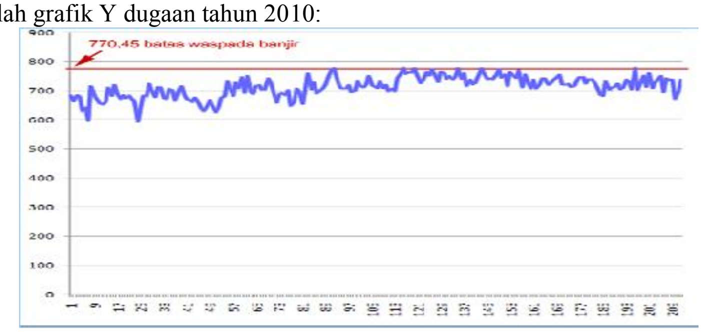
Y_t : $1092.7828 + 0.9291CH_{t-6} - 24.4848T_{t-2} - 0.06245PM_{t-2} + 1.4706KB_{t-2}$

Dari persamaan didapat bahwa pengaruh paling besar diberikan oleh temperatur sisanya mempunyai pengaruh dibawah angka dua. Curah hujan dan kelembaban nisbi memberi pengaruh positif yang berarti semakin tinggi curah hujan dan kelembaban nisbi maka akan semakin tinggi pula ketinggian permukaan airnya. Sebaliknya temperatur dan penyinaran matahari menunjukan pengaruh negatif.

Model Y_t : $1092.7828 + 0.9291CH_{t-6} - 24.4848T_{t-2} - 0.06245PM_{t-2} + 1.4706KB_{t-2}$ dapat kita gunakan untuk meramalkan maksimal kejadian dua hari sesudahnya. Dan sebagai indikator banjir penulis menggunakan rumus $\bar{u} \pm sd$. Rumus ini adalah rumus statistik kualitas kontrol (*quality control*) yang mengambil rata-rata sebagai nilai tengah dan standart deviasi sebagai batas atas dan batas bawah.

Dengan subtitusi nilai rata-rata dan nilai standart deviasi dari analisis deskriptif didapat rumus tersebut menjadi 713,94 ± 56,51. Yang berarti bawah jika data yang diramalkan melebihi batas atas yaitu 713,94 + 56,51 atau 770,45. Maka status akan dikatakan sebagai status waspada karena kemungkinan besar ketinggian permukaan air di Manggarai akan menyebabkan banjir.

Berikut adalah grafik Y dugaan tahun 2010:



Gambar 6. Grafik Y Dugaaan

5. KESIMPULAN

Berdasarkan hasil analisis dan pembahasan seperti yang telah di uraikan, maka dapat disimpulkan bahwa Model yang dihasilkan adalah Y_t : $1092.7828 + 0.9291CH_{t-6} - 24.4848T_{t-2} - 0.06245PM_{t-2} + 1.4706KB_{t-2.}$

Indikator data klimatologi yang mempengaruhi ketinggian permukaan air adalah Temperatur dengan ketinggian permukaan air merupakan variabel yang memiliki hubungan yang paling kuat. Variabel ini memiliki hubungan secara negatif yang berarti ketika temperatur turun maka nilai ketinggian permukaan air akan naik. Sedang koefisien determinasi memiliki nilai sebesar $R^2 = 0.4056$ dan statistik Durbin Watson sebesar DW = 0.7429. keduanya berguna untuk mengetahui keakuratan model.

UCAPAN TERIMAKASIH

Dalam kesempatan ini penulis mengucapkan terima kasih kepada **Henry Eka Saputra** alumni jurusan Statistika FST Binus University yang telah membantu penulis dalam pelaksanaan penelitian ini sehingga dapat diselesaikan sesuai dengan rencana.

DAFTAR PUSTAKA

Dunteman, 1985, *Introduction to Multivariate Analysis*, SAGE Publication, USA

Hair, Anderson, 1998, *Multivariate Data Analysis*, 5 ed, Prentice Hall, Newyork

Markidakis, 1999, *Metode dan Aplikasi Peramalan*, Erlangga, Jakarta

Netter J, Kutner MH Nachtseim, Wasseman, 1996, *Applied Linear Regression Models*, 3ed, McGrawHill, Newyork

Shneiderman, 1998, *Designing the user Interface Strategies for Effective HCI*, 3ed, Addison Wesley, England